Stochastic Bandwidth Estimation in Networks With Random Service

Ralf Lübben, *Student Member, IEEE*, Markus Fidler, *Senior Member, IEEE*, and Jörg Liebeherr, *Fellow, IEEE*

***Abstract*—Numerous methods for available bandwidth estimation have been developed for wireline networks, and their effectiveness is well-documented. However, most methods fail to predict bandwidth availability reliably in a wireless setting. It is accepted that the increased variability of wireless channel conditions makes bandwidth estimation more difficult. However, a (satisfactory) explanation why these methods are failing is missing. This paper seeks to provide insights into the problem of bandwidth estimation in wireless networks or, more broadly, in networks with random service. We express bandwidth availability in terms of bounding functions with a defined violation probability. Exploiting properties of a stochastic min-plus linear system theory, the task of bandwidth estimation is formulated as inferring an unknown bounding function from measurements of probing traffic. We present derivations showing that simply using the expected value of the available bandwidth in networks with random service leads to a systematic overestimation of the traffic departures. Furthermore, we show that in a multihop setting with random service at each node, available bandwidth estimates requires observations over (in principle infinitely) long time periods. We propose a new estimation method for random service that is based on iterative constant-rate probes that take advantage of statistical methods. We show how our estimation method can be realized to achieve both good accuracy and confidence levels. We evaluate our method for wired single-and multihop networks, as well as for wireless networks.**

***Index Terms*—Bandwidth estimation, communications technology, computer networks, IP networks, random systems, stochastic network calculus.**

I. Introduction

THE OBJECTIVE of available bandwidth estimation is to infer the service offered by a network path from traffic measurements taken at end-systems only. In bandwidth estimation methods, end-systems exchange timestamped probe packets and study the dispersion of these packets after they have traversed a network of nodes. In recent years, available bandwidth estimation has attracted significant interest, and a wide variety of measurement tools and techniques have been developed, e.g., [16]–[18], [34], [39], and [42]. Many of the most popular methods for available bandwidth estimation are based on congestion-inducing packet trains, where a packet train consists of a sequence of probe packets. By sending packet trains at a rate exceeding the available bandwidth, the network becomes congested, thereby imprinting information on the network state on the dispersion of probe packets.

Virtually all available bandwidth methods were developed for wireline networks, where communication channels consist of fixed-capacity links, and where the available bandwidth of a link is given by its unconsumed capacity. Some of these methods have been adapted for wireless networks (see Section II), in particular WiFi networks, however, they generally lack the robustness and reliability achieved in fixed-capacity wireline environments. A potential source of errors are unsuitable model assumptions. For example, many methods interpret the packet dispersion under the assumption that probe traffic flows through one or more fixed-capacity first-in–first-out (FIFO) links that experience cross traffic [16], [34], [42]. Even though FIFO queueing may be highly prevalent in wired network infrastructures today, FIFO assumptions are difficult to justify in wireless multiaccess networks [5], [6], [38]. Another widely used model assumption is that networks are work-conserving, i.e., they generate output whenever traffic is ready for transmission. However, as pointed out in [23], latencies incurred during channel access lead to non-work-conserving systems.

In this paper, we investigate fundamental difficulties of measuring the available bandwidth in wireless networks with congestion-inducing packet trains. Rather than revising or adapting wireline approaches to wireless channels, e.g., by trying to eliminate superimposed random "noise," we seek to develop from the ground up a new modeling and inference approach for networks that are subject to randomness of both traffic and transmission channels. We dispense with the modeling assumption of a work-conserving queueing system and, taking advantage of concepts from the stochastic network calculus [20], replace it with that of a general stationary system.

The point of departure of our efforts is a recent system-theoretic approach of bandwidth estimation [28]. Here, the network is viewed as a time-invariant deterministic system where throughput and delays of traffic are governed by an unknown bounding function, referred to as *service curve*. Service curves can express work-conserving as well as non-work-conserving systems. For example, the function $\mathcal{S}(t) = R\max\{0, t-T\}$ expresses a latency-rate service curve of a non-work-conserving system with a latency of T and rate R. It can be argued

Manuscript received April 03, 2012; revised December 29, 2012; accepted March 03, 2013; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor L. Andrew. Date of publication May 21, 2013; date of current version April 14, 2014. The work of R. Lübben and M. Fidler was supported in part by the German Research Foundation (DFG) under an Emmy Noether Grant. The work of J. Liebeherr was supported in part by an NSERC Strategic Project. This work is an extended version of a paper that appeared in the Proceedings of the IEEE International Conference on Computer Communications (INFOCOM), Shanghai, China, April 10–15, 2011.

R. Lübben and M. Fidler are with the Department of Electrical Engineering and Computer Science, Leibniz Universität Hannover, Hannover 30167, Germany.

J. Liebeherr is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: jorg@comm.utoronto.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2013.2261914

that existing congestion inducing probing methods assume that a network is a time-invariant deterministic system. As a case in point, packet train probing schemes similar to Pathload [17] and Pathchirp [39] were shown to be aligned with a system-theoretic interpretation in that they can reliably extract the shape of convex service curves of a time-invariant deterministic system. A strength of the system-theoretic approach is that service curves offer a natural extension to networks with several bottleneck links by exploiting properties of the network calculus [8], [26].

A limitation of the system-theoretic approach from [28] is the assumption that the measured system is time-invariant deterministic, meaning that identical packet probes, e.g., packet pairs or packet trains, sent at two different time instances experience the same backlog or delay. Even in networks with random traffic load or random link capacities, this can be an appropriate model as long as the timescale of measurements is small compared to the timescale at which resource availability in the network changes. In such cases, the network can be viewed as a state-dependent deterministic system, where a single sample of the available bandwidth can be interpreted as being conditioned on the current state of the network. Evaluating a large number of samples corresponds to computing a conditional average of the system state. However, when network or traffic characteristics change on a short timescale, e.g., during individual measurements, a time-invariant and deterministic system characterization is not suitable.

The main contribution of this paper is the development of a system-theoretic approach to bandwidth estimation of systems with random service, which can account for variability due to statistical properties of network traffic and transmission channels on short timescales. We develop the foundations for the approach (in Section III) by first expressing random service in a network by ε-effective service curves from the stochastic network calculus [7]. Working within the framework of the network calculus, we show that it is feasible to compute an ε-effective service curve of a network from packet trains with fixed inter packet gap in each train [16], [17], [34]. By relating the service curves of the network calculus to a common definition of available bandwidth, we can express limitations of existing probing methods in a stochastic stationary system. For example, for a single-link network the service curve describing the unused capacity on a link coincides with the commonly used definition of available bandwidth. In contrast, for multilink networks, we show that the definition of end-to-end available bandwidth can be recovered only as a time limit, and that the available bandwidth may overestimate the usable service over short timescales. Moreover, in a system with random service, the expected value of the available bandwidth provides only an optimistic estimate of the traffic departures.

Equipped with a system theory for bandwidth estimation of stochastic systems, we address the development of a practical probing scheme (in Section IV). For example, a challenge of using packet train probing in a stochastic system is that packet trains may push the system into a nonstationary state. In principle, observing the system in a stationary state requires packet trains with infinite length. We show that, in practice, it is possible to detect stationarity with finite packet trains using stationarity tests and other statistical methods and to dynamically adapt packet trains to the required length. We also study the required lengths of packet trains and the required number of repeated measurements.

Using measurement results from a controlled testbed, we quantify the effect of variability on the estimated service and observe the impact of the burstiness of cross traffic, access delays, and retransmissions on service availability. We present measurement results for wired single-hop and multihop networks as well as for wireless networks.

An implication of our study is that the widely used assumption of work-conserving fixed-capacity FIFO links with cross traffic can be replaced by a more general network model without specific requirements on the multiplexing method. This may open the field of bandwidth estimation to network environments where FIFO or work-conserving assumptions are not justified.

The remainder of this paper is structured as follows. In Section II, we discuss related work on approaches to bandwidth estimation, with a focus on congestion-inducing methods. In Section III, we derive a stochastic min-plus approach for estimating networks with random service. In Section IV, we consider practical aspects of bandwidth estimation of a stochastic system. In Section V, we provide an experimental validation of our method. Section VI gives brief conclusions.

II. Available Bandwidth Estimation

The term *available bandwidth* denotes the capacity that is left unused by other traffic in the network, referred to as cross traffic. For a link h, it can be expressed for any time interval $(\tau, t]$ as [29]

$$\alpha^h(\tau,t) = \frac{1}{t-\tau}\int_{\tau}^{t} C^h(x)\left(1-u^h(x)\right)dx \tag{1}$$

where $C^h(t)$ is the possibly time-varying capacity of the link and $u^h(t) \in [0,1]$ is its utilization by cross traffic at time t. For cross traffic with a long-term average rate λ^h, the limit $\liminf_{\delta\to\infty} \alpha^h(\tau,\tau+\delta) = C^h - \lambda^h$ is referred to as *long-term available bandwidth*. The end-to-end available bandwidth of a network path is frequently defined as the minimum of the available bandwidths of all traversed links $(h = 1,\ldots,H)$ [29], [30]

$$\alpha^{\text{net}}(\tau,t) = \min_h \left\{\alpha^h(\tau,t)\right\}. \tag{2}$$

A. Bandwidth Estimation of Work-Conserving Systems

Many bandwidth estimation tools assume a fluid time-invariant network model with work-conserving FIFO scheduling, where the relation between the incoming rate r_{I} and the outgoing rate r_{O} of a constant bit rate (CBR) probe at a link, referred to as rate response curve [29], [35], is given by

$$\frac{r_{\text{I}}}{r_{\text{O}}} = \begin{cases} 1, & r_{\text{I}} \le C-\lambda \\ \frac{r_{\text{I}}+\lambda}{C}, & r_{\text{I}} > C-\lambda. \end{cases} \tag{3}$$

Random cross traffic is often interpreted as distorting the traffic dispersion given by response curves. To eliminate the random distortions, estimation tools have applied averaging [16], [42], linear regression [34], and Kalman filtering [11]. For packet pair probes, the corresponding function describing the dispersion of the probes, is called gap response curve [16], [29]. Several works have extended the deterministic CBR traffic model

for the response curve at a FIFO system to stochastic ones, e.g., [10], [14], [29], [30], [36], and [37]. A queueing theoretic framework for bandwidth estimation is analyzed in [29], where it is shown that the assumption of fluid CBR traffic generates an upper bound for the available bandwidth, and that the deviation can be resolved using packet trains of infinite length. The work is extended to multihop networks in [30]. In [10], a distribution for the output gap is derived for a general arrival process in conjunction with parameter estimation for known cross-traffic distributions. Fundamental limitations of active probing are analyzed in [33] based on a queueing model of a FIFO system.

Methods that do not explicitly assume FIFO scheduling but are compatible with this assumption are, e.g., Pathchirp [39] and Pathload [17], [18]. Both methods increase their probing rate until an increase of one-way delays or, equivalently, queueing delays of probe packets is detected. Pathload specifies the available bandwidth as a range, to capture its time-varying nature. The underlying deterministic network model is relaxed to filter out short-term fluctuations in the detection of long-term trends.

B. Bandwidth Estimation in Wireless Networks

Bandwidth estimation in wireless networks is more difficult for a variety of reasons. Link capacities in wireless networks are not constant due to interference and time-varying channel conditions. Also, wireless networks may not behave like FIFO systems. For example, a measurement study of IEEE 802.11 networks [6] showed that the multiplexing of the distributed coordination function (DCF) is similar to a fair queueing algorithm. Furthermore, inherent delays due to media access, retransmissions, and other factors create a non-work-conserving system [23], where the transmission channel can be idle even if traffic is pending. Proposed approaches in the literature have addressed these issues by adapting or revising bandwidth estimation methods developed for wireline networks.

Time-varying link capacities are addressed in [27]. Instead of estimating the link capacity by the smallest gap between two successive probe packets, the median of a sequence of probe gaps is used. In [22], the authors observe dependencies of the cross traffic rate and the probe packet size on the capacity of IEEE 802.11 networks. A passive measurement approach to estimate the available bandwidth in IEEE 802.11 networks that does not assume FIFO scheduling is presented in [24]. Non-FIFO scheduling is also investigated in [5], and [38] where the rate response curve is adapted to the channel access in IEEE 802.11 networks and to the wireless channel. Additionally, the transient behavior of access delays and its impact on the response curve are described in [38]. Kalman filters are adapted to the channel access and to the wireless channel in IEEE 802.11 networks in [6] and [21], respectively. Recently, attention has been given to measuring the service in wireless home networks. In addition to the rate of the available bandwidth, estimation approaches for home networks in [9] address access delays as important metrics. In [23], correlations of access delay measurements are used to differentiate, among others, traffic congestion and hidden terminal problems. The study emphasizes the need for non-work-conserving models due to the significant role of non-queueing-related delay components.

C. Bandwidth Estimation of Min-Plus Linear Systems

System theory offers an alternative model for bandwidth estimation methods [1], [15], [28]. A network is represented as a general system with traffic arrivals $A(t)$ as input and traffic departures $D(t)$ as output of the system. $A(t)$ and $D(t)$ are nonnegative and nondecreasing functions that denote the cumulative number of bits seen in an interval $(0,t]$. By convention, $A(0) = 0$, and due to causality, $D(t) \leq A(t)$ for all $t \geq 0$. We sometimes use the shorthand notation $A(\tau,t) = A(t) - A(\tau)$ to denote the arrivals in the interval $(\tau,t]$. We use $B(t) = A(t) - D(t)$ to describe the backlog of the system at time t and the smallest number d such that $D(t+d) \geq A(t)$ to characterize the delay at time t.

A system is called time-invariant if a time-shifted arrival function $A(t-\tau)$ results in a time-shifted departure function $D(t-\tau)$. Consider two arrival functions A_1, A_2 and their corresponding departure functions D_1, D_2 of a system. A system is called min-plus linear when arrivals of the form $\min\{A_1(t), A_2(t)\} + c$, where $c \geq 0$ is constant, result in departures $\min\{D_1(t), D_2(t)\} + c$. The form of the arrivals can be viewed as a linear combination in an algebra, where the minimum replaces the usual addition, and the addition takes the role of the multiplication. This gives rise to the min-plus algebra formulation of the network calculus.

The behavior of any time-invariant min-plus linear system can be characterized by a nonnegative and nondecreasing function $\mathcal{S}(t)$, referred to as *service curve*. The service curve relates departures and arrivals of a system by

$$D(t) = \inf_{\tau \in [0,t]} \{A(\tau) + \mathcal{S}(t-\tau)\} =: A \otimes \mathcal{S}(t) \qquad (4)$$

where the operator $\otimes$ is the convolution under the min-plus algebra. In the language of system theory, the service curve is the impulse response of a min-plus linear system [26]. Service curves can be used to describe a wide range of systems. A work-conserving constant-rate link is expressed by $\mathcal{S}(t) = Rt$ with $(R > 0)$, and a non-work-conserving pure-delay server that imposes a latency $d > 0$ is given by $\mathcal{S}(t) = \infty$ if $t \geq d$ and $\mathcal{S}(t) = 0$ if $t < d$.

For min-plus linear systems, bandwidth estimation can be expressed as the inversion problem of obtaining $\mathcal{S}$ from $D = A \otimes \mathcal{S}$ [28], where A and D are arrival and departure functions of probing traffic. A solution to the inversion problem can be obtained from constant-rate packet train arrivals $A(t) = rt$ and measurements of the departure traffic $D(t)$. Using the maximum system backlog, expressed as $B_{\max}(r) = \sup_\tau \{A(\tau) - D(\tau)\}$, the service curve can be computed as

$$\mathcal{S}(t) = \sup_{r \geq 0} \{rt - B_{\max}(r)\}.$$

An advantage of a system-theoretic view is a straightforward extension to multihop settings. Given a network of H systems in sequence where $\mathcal{S}^h (h = 1, \ldots, H)$ denotes the service curve of the hth system, a service curve $\mathcal{S}^{\text{net}}$ for the entire sequence of H systems is given by the min-plus convolution $\mathcal{S}^{\text{net}} = \mathcal{S}^1 \otimes \ldots \otimes \mathcal{S}^H$. The equivalent replacement of a sequence of single-hop systems by a single system with service curve $\mathcal{S}^{\text{net}}$ provides a handle on an analysis of multihop networks.

While (4) is suitable for expressing the service offered by constant-rate links, traffic regulators, or fair schedulers, it assumes

linearity under the min-plus algebra since the convolution is a linear operation. For nonlinear systems, (4) can be relaxed to provide a linear lower bound of the form $D \geq A \otimes \mathcal{S}$ [26]. Notable nonlinear systems that are used in many actual networks are FIFO schedulers [13], where the service obtained by a flow in an overloaded system depends on its arrival rate; see (3). In [28], it is argued that networks can be viewed as linear systems that transition to a nonlinear regime when the network becomes saturated, and it is shown that the transition can be observed using suitable nonlinearity tests.

The main limitation of the system-theoretic approach to bandwidth estimation is that the measured system must satisfy time invariance, i.e., the service curve in (4) is a deterministic function that does not depend on the absolute values of τ and t, but only on the length of the time interval $\delta = t-\tau$. Dispensing with this assumption requires a system-theoretic framework where the network service is a random process.

III. Inference of a Random Service

In this section, we develop the foundation for a stochastic bandwidth estimation methodology for networks with stationary random service. Starting from a system characterization by random service processes as defined in [8], we link the service processes to the stochastic network calculus [20] and also to the available bandwidth as given in (1). In detail, we proceed as follows. In Section III-A, we show that stochastic bounds of random service processes can be expressed in terms of ε-effective service curves $\mathcal{S}^\varepsilon$ [7] from the stochastic network calculus, which enables us to express statistical performance bounds of the type $\mathsf{P}[B > x] \leq \varepsilon$. We show how to infer an ε-effective service curve from measurements of probing traffic in Section III-B. In Section III-C, we address the relation of effective service curves to the available bandwidth given by (1) and (2) and explore conditions where the available bandwidth systematically overestimates the actual departures of a network. In Section III-D, we show that by expressing arrivals and departures using a max-plus algebra, the inference of service can be done directly from packet timestamps. This last step will be used in the next section, where we develop a practical probing methodology.

A. Systems With Random Service in the Network Calculus

Given a system with random service, the assumption of time invariance, on which the definition of the service curve $\mathcal{S}(t)$ in (4) is based, does not hold. To this end, we substitute $\mathcal{S}(t)$ by a bivariate random process $S(\tau,t)$, which expresses a random service experienced in the time interval $(\tau,t]$. We refer to $S(\tau,t)$ as a *service process*. Similarly as in (4), the departures of a min-plus linear random system can be related to its arrivals by the service process as [8]

$$D(t) = \inf_{\tau\in[0,t]} \{A(\tau) + S(\tau,t)\} =: A \otimes S(t). \tag{5}$$

In the stochastic network calculus, random service can be expressed by ε-effective service curves, which express a non-random time-invariant bound on the service that can be violated with probability ε. An ε-effective service curve $\mathcal{S}^\varepsilon(t)$ specifies a service guarantee of the form [7]

$$\mathsf{P}\left[D(t) \geq \inf_{\tau\in[0,t]} \{A(\tau) + \mathcal{S}^\varepsilon(t-\tau)\}\right] \geq 1-\varepsilon. \tag{6}$$

The following lemma links the definitions in (5) and (6), in that it specifies ε-effective service curves as a stationary bound of a random service process.

Lemma 1: Given a system with service process $S(\tau,t)$ as in (5). Any function $\mathcal{S}^\varepsilon(t)$ that satisfies the sample path bound

$$\mathsf{P}\left[S(\tau,t) \geq \mathcal{S}^\varepsilon(t-\tau), \quad \forall\tau\right] \geq 1-\varepsilon$$

for $t \geq 0$ is an ε-effective service curve in the sense of (6) of the system.

The advantage of the ε-effective service curve $\mathcal{S}^\varepsilon(t-\tau)$ compared to the service process $S(\tau,t)$ is that it is time-invariant, i.e., $\mathcal{S}^\varepsilon(t-\tau)$ depends on the duration $t-\tau$, but it does not depend on the location of the time interval $(\tau,t]$. Later, we will relate the service process (5) to the available bandwidth definition in (1).

Proof: Consider a sample path $S^\omega(\tau,t)$ of $S(\tau,t)$ and fix $t \geq 0$. If $S^\omega(\tau,t) \geq \mathcal{S}^\varepsilon(t-\tau)$ for all $\tau \in [0,t]$, it follows from the monotonicity of the min-plus convolution that

$$D(t) = A \otimes S^\omega(t) \geq A \otimes \mathcal{S}^\varepsilon(t).$$

Since, by assumption, the condition $S^\omega(\tau,t) \geq \mathcal{S}^\varepsilon(t-\tau)$ holds for all $\tau \in [0,t]$ with probability at least $1-\varepsilon$, the claim is proven. ∎

B. Estimation of Effective Service Curves

We next show how an ε-effective service curve can be obtained from constant-rate packet train probes of the form $A(t) = rt$. To this end, we phrase the backlog as a function of the probing rate, that is, $B(r,t) = rt - D(t)$. The steady-state backlog for $t \to \infty$ is abbreviated by $B(r)$. We define the quantile of the backlog distribution $B^\xi(r,t)$ as

$$B^\xi(r,t) = \inf\{x \geq 0 : \mathsf{P}[B(r,t) \leq x] \geq 1-\xi\} \tag{7}$$

where ξ denotes a violation probability. With this definition, Theorem 1 provides the foundation for a packet train based estimation method.

Theorem 1: Given a system with service process $S(\tau,t)$ as in (5). Select a finite set R of rates $r \geq 0$. For all $t \geq 0$, the function

$$\mathcal{S}^\varepsilon(t) = \max_{r\in R}\left\{rt - B^\xi(r)\right\}$$

is an ε-effective service curve in the sense of (6) of the system with violation probability $\varepsilon = \sum_{r\in R}\xi$.

Proof: From $B(t) = A(t) - D(t)$ and (5), it follows that

$$B(t) = \sup_{\tau\in[0,t]} \{A(\tau,t) - S(\tau,t)\}. \tag{8}$$

The supremum in (8) implies that $B(t) \geq A(\tau,t) - S(\tau,t)$ for all $\tau \in [0,t]$, permitting us to write

$$S(\tau,t) \geq A(\tau,t) - B(t) \qquad \forall \tau \in [0,t].$$

Inserting $A(\tau,t) = r(t-\tau)$ and using the backlog quantile yields

$$\mathsf{P}\left[S(\tau,t) \geq r(t-\tau) - B^{\xi}(r,t),\ \forall \tau\right] \geq 1-\xi.$$

Using the complement and applying the union bound for a set of rates R, it follows that

$$\mathsf{P}\left[\bigcup_{r\in R}\left\{S(\tau,t) < r(t-\tau) - B^{\xi}(r,t),\ \forall \tau\right\}\right] \leq \sum_{r\in R}\xi$$

and hence

$$\mathsf{P}\left[S(\tau,t) \geq \max_{r\in R}\left\{r(t-\tau) - B^{\xi}(r,t)\right\},\ \forall \tau\right] \geq 1 - \sum_{r\in R}\xi.$$

With Lemma 1, we obtain that $\mathcal{S}^{\varepsilon}(t-\tau)$ defined as

$$\mathcal{S}^{\varepsilon}(t-\tau) = \max_{r\in R}\left\{r(t-\tau) - B^{\xi}(r,t)\right\}$$

for all $\tau \in [0,t]$ is an ε-effective service curve with violation probability $\varepsilon = \sum_{r\in R}\xi$. Letting $t \to \infty$ and inserting the steady-state backlog quantile $B^{\xi}(r)$ completes the proof. ∎

Theorem 1 gives rise to a method for service curve estimation for networks with random service using packet train probes sent at different rates r. Each packet train provides a sample of the steady-state backlog $B(r)$. Taking many samples provides an empirical distribution of the steady-state backlog from which the quantile, denoted by $\widetilde{B}^{\xi}(r)$ and given in (7), is obtained. Applying Theorem 1, we can compute an estimate of an ε-effective service curve as $\widetilde{\mathcal{S}}^{\varepsilon}(t) = \max_{r\in R}\{rt - \widetilde{B}^{\xi}(r)\}$. The choice of the probing set R presents a tradeoff. The accuracy of the estimate increases by adding probing rates. At the same time, adding probing rates increases the violation probability due to the use of the union bound in Theorem 1.

Since our method uses steady-state backlogs $B(r)$ to obtain service curve estimates, the convergence to a steady state is vital. The following lemma is important as it provides existence of the steady-state backlog distribution as long as the probing rate does not exceed the limiting rate of the service process, defined as $\liminf_{t\to\infty}(S(\tau,t)/(t-\tau))$.

Lemma 2: Given arrivals $A(\tau,t)$ at a system with service process $S(\tau,t)$, satisfying (5), where $A(\tau,t)$ and $S(\tau,t)$ are jointly stationary in the strong sense[1].

1) The backlog $B(t)$ is stochastically increasing in t.
2) If for all t, it holds that

$$\limsup_{\delta\to\infty}\frac{A(t-\delta,t)}{\delta} < \liminf_{\delta\to\infty}\frac{S(t-\delta,t)}{\delta}$$

almost surely, the backlog converges in distribution to a finite random variable B.

Note that we extended the processes $A(\tau,t)$ and $S(\tau,t)$ from $0 \leq \tau \leq t < \infty$ to $-\infty < \tau \leq t < \infty$. The lemma generalizes [8, Lemma 9.1.4] from a constant-rate server to a server with random service. The proof closely follows [8].

Proof: From (8), it follows for any x and $\vartheta > 0$ that

$$\begin{aligned}\mathsf{P}\left[B(t+\vartheta) \geq x\right] &= \mathsf{P}\left[\sup_{\tau\in[0,t+\vartheta]}\left\{A(\tau,t+\vartheta) - S(\tau,t+\vartheta)\right\} \geq x\right] \\ &\geq \mathsf{P}\left[\sup_{\tau\in[0,t]}\left\{A(\tau+\vartheta,t+\vartheta) - S(\tau+\vartheta,t+\vartheta)\right\} \geq x\right].\end{aligned}$$

From the assumption of joint stationarity, $A(\tau+\vartheta,t+\vartheta) - S(\tau+\vartheta,t+\vartheta)$ is equal in distribution to $A(\tau,t) - S(\tau,t)$ for all τ, t, and $\vartheta > 0$. The last line equals $\mathsf{P}[B(t) \geq x]$ so that we get $\mathsf{P}[B(t+\vartheta) \geq x] \geq \mathsf{P}[B(t) \geq x]$, thus proving the first claim.

For the second claim, if the given inequality holds, then there exists a finite random variable

$$T = \sup\left\{\delta \geq 0 : A(t-\delta,t) \geq S(t-\delta,t)\right\}$$

for any t. Consequently, $A(t-\delta,t) < S(t-\delta,t)$ holds for all $\delta > T$ with probability one. Moreover, since $A(t-\delta,t)$ is nondecreasing in $\delta \geq 0$, we have $A(t-\delta,t) < S(t-T-\vartheta,t)$ for all $0 \leq \delta \leq T$ and any $\vartheta > 0$. Combining the two statements and using that $S(t-\delta,t)$ for $\delta \geq 0$ and $S(t-T-\vartheta,t)$ are nonnegative yields

$$A(t-\delta,t) - S(t-\delta,t) \leq S(t-T-\vartheta,t)$$

for all $\delta \geq 0$. Hence

$$\sup_{\delta\geq 0}\left\{A(t-\delta,t) - S(t-\delta,t)\right\} \leq S(t-T-\vartheta,t).$$

With $\sup_{\delta\geq 0}\{A(t-\delta,t) - S(t-\delta,t)\} = B(t)$ from (8), it follows for any x that

$$\sup_{t}\left\{\mathsf{P}\left[B(t) \geq x\right]\right\} \leq \sup_{t}\left\{\mathsf{P}\left[S(t-T-\vartheta,t) \geq x\right]\right\}.$$

Since T is finite and $B(t)$ is stochastically increasing, there exists a finite random variable B such that

$$\lim_{t\to\infty}\mathsf{P}\left[B(t) \geq x\right] = \sup_{t}\left\{\mathsf{P}\left[B(t) \geq x\right]\right\} = \mathsf{P}[B \geq x]$$

completing the proof of the second claim. ∎

Example: We illustrate our estimation approach and its achievable accuracy using a computed numerical example. We consider a discrete-time system consisting of a random non-work-conserving On–Off server. In each time-slot, the server performs an independent Bernoulli trial, where it forwards one unit of data with probability $p = 0.1$. This system is non-work-conserving since, even if there is a backlog, the server remains idle with probability of $1-p$. The system is probed with constant-rate probes that are each sent for a duration of 1000 time-slots. Fig. 1 shows the estimated ε-effective service curve with $\varepsilon = 10^{-3}$ computed with Theorem 1. As indicated by the thin dash-dotted lines, each probing rate r contributes a linear segment with slope r displaced by $-B^{\xi}(r)$ to the service curve estimate. Analytical upper and lower bounds of the service curve are included in the graph as a reference. The upper bound is computed as the amount of data forwarded

[1]Stationarity in the strong sense means that $\mathsf{P}[A(\tau,t) \leq x] = \mathsf{P}[A(\tau+\vartheta,t+\vartheta) \leq x]$ and joint stationarity $\mathsf{P}[A(\tau,t) \leq x, S(\tau,t) \leq y] = \mathsf{P}[A(\tau+\vartheta,t+\vartheta) \leq x, S(\tau+\vartheta,t+\vartheta) \leq y]$ for all $\tau \leq t$ and all ϑ.

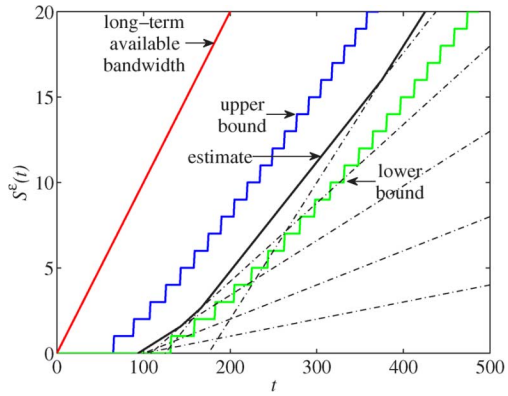

Fig. 1. Service curves of an On–Off server. The service curve estimate is composed of linear segments that are each obtained by a single probing rate. Analytical lower and upper bounds are included for comparison. The limiting rate of the service curve estimate converges to the long-term available bandwidth. Compared to the long-term available bandwidth, the service curve provides significant details on the timescales of service availability.

in t time-slots with probability at least $1-\varepsilon$. By assumption, this can be computed from a binomial distribution with parameter p. Due to the discrete model, the bound is a staircase function. The lower bound is computed from the upper bound by an application of the union bound. The depicted service curve estimate reflects a fluid flow traffic model, which is different from the staircase function obtained from the binomial distribution. Discrete-sized data packets will be incorporated into our method in Section III-D. For comparison, we show the service curve with a constant rate computed as the average of (1). The non-work-conserving aspect of the system is captured by the initial latency of about 100 time-slots, where the estimate of $\mathcal{S}^\varepsilon(t)$ in Fig. 1 evaluates to zero. While the limiting rate of the service curve estimate, defined as $\lim_{t\to\infty} \tilde{\mathcal{S}}^\varepsilon(t)/t$, is equal to the long-term available bandwidth, neglecting the latency inherent to the system, the long-term available bandwidth overestimates the service in the system (by violating the upper bound). This illustrates that the ε-effective service curve is a nontrivial generalization of the available bandwidth concept since the service curve can capture that the available service rate is variable at different timescales.

C. Connection to Available Bandwidth

In this section, we relate our stochastic system-theoretic approach to the available bandwidth definition from (1). We will refer to the traffic in the system as cross traffic, and denote arrivals and departures by A_c and D_c, respectively. Probe traffic arrivals and departures are denoted by A_p and D_p, respectively. We will show that the service process of a work-conserving system is equivalent to the available bandwidth from (1). At the same time, we show that the commonly used probabilistic version, namely, the expected available bandwidth, necessarily leads to an overestimation of the expected departures of a system. Also, we show that a widely used definition of available bandwidth for multihop systems $\alpha^\mathrm{net}(\tau,\tau+\delta)$, given in (2), can be recovered only in the limit $\delta\to\infty$.

1) Single-Hop Systems: Given a work-conserving server with service process $S(\tau,t)$, where $S(\tau,t)$ denotes the amount of service available in $(\tau,t]$ [8]. For any $t\geq 0$, let $\tau\geq 0$ denote the beginning of the last busy period before t. From the work-conserving property, it holds that $D(t)=D(\tau)+S(\tau,t)$. Now, let $D(t)=D_\mathrm{c}(t)+D_\mathrm{p}(t)$ be composed of cross traffic and probe traffic, respectively. It follows that

$$D_\mathrm{p}(t)=D_\mathrm{p}(\tau)+\underbrace{S(\tau,t)-D_\mathrm{c}(\tau,t)}_{=:S_l(\tau,t)} \tag{9}$$

where the service process $S_l(\tau,t)$ denotes the service left over by cross traffic in $(\tau,t]$. By choice of τ, it holds that $D_\mathrm{p}(\tau)=A_\mathrm{p}(\tau)$ so that $D_\mathrm{p}(t)=A_\mathrm{p}(\tau)+S_l(\tau,t)$. Given that there exists at least one $\tau\in[0,t]$ such that $D_\mathrm{p}(t)=A_\mathrm{p}(\tau)+S_l(\tau,t)$, we conclude that $D_\mathrm{p}(t)\geq\inf_{\tau\in[0,t]}\{A_\mathrm{p}(\tau)+S_l(\tau,t)\}=A_\mathrm{p}\otimes S_l(t)$. Note that since the min-plus convolution evaluates all $\tau\in[0,t]$, it is not confined to the beginning of a specific busy period. Also, from (9), it follows for all $\tau\in[0,t]$ that $D_\mathrm{p}(t)\leq A_\mathrm{p}(\tau)+S_l(\tau,t)$ since $D_\mathrm{p}(\tau)\leq A_\mathrm{p}(\tau)$ due to causality. It follows that $D_\mathrm{p}(t)\leq\inf_{\tau\in[0,t]}\{A_\mathrm{p}(\tau)+S_l(\tau,t)\}$. Combining the lower and the upper bound, we obtain $D_\mathrm{p}(t)=A_\mathrm{p}\otimes S_l(t)$, that is, $S_l(\tau,t)$ satisfies (5).

Next, we prove that the available bandwidth defined in (1) coincides with the leftover service process $S_l(\tau,t)$. Consider the stochastic leftover service process of a constant-rate link $S(\tau,t)=C(t-\tau)$ that is utilized by random cross traffic with intensity $u(t)$. We obtain with (1) that

$$\alpha(\tau,t)=\frac{\int_\tau^t C\,(1-u(x))\,dx}{t-\tau}=\frac{S(\tau,t)-D_\mathrm{c}(\tau,t)}{t-\tau}$$

since the amount of cross traffic departures is determined by the utilization in $(\tau,t]$ as $D_\mathrm{c}(\tau,t)=C\int_\tau^t u(x)dx$. It follows that

$$\alpha(\tau,t)=\frac{S_l(\tau,t)}{t-\tau}. \tag{10}$$

Thus, the available bandwidth from (1) for a specific time interval corresponds to the rate of the leftover service process in this time interval.

In bandwidth estimation, the time-varying nature of the available bandwidth is frequently covered by using the expected value. However, while $\alpha(\tau,t)$ can be related to the leftover service process, the following lemma states that $\mathsf{E}[\alpha(\tau,t)]$ systematically overestimates the departures.

Lemma 3: Given $D_\mathrm{p}(t)=A_\mathrm{p}\otimes S_l(t)$. It holds that

$$\mathsf{E}\left[D_\mathrm{p}(t)\right]\leq A_\mathrm{p}\otimes\mathsf{E}[S_l](t).$$

Constructing examples where Lemma 3 does not hold with equality is straightforward.

Proof: Taking expectations we have

$$\begin{aligned}\mathsf{E}\left[D_\mathrm{p}(t)\right]&=\mathsf{E}\left[\inf_{\tau\in[0,t]}\{A_\mathrm{p}(\tau)+S_l(\tau,t)\}\right]\\&=\sum_{\omega\in\Omega}p^\omega\inf_{\tau\in[0,t]}\{A_\mathrm{p}(\tau)+S_l^\omega(\tau,t)\}\end{aligned}$$

where S_l^Ω is the sample space containing sample paths S_l^ω that occur with probability p^ω each. For any choice of $\tau'\in[0,t]$

$$p^\omega\inf_{\tau\in[0,t]}\{A_\mathrm{p}(\tau)+S_l^\omega(\tau,t)\}\leq p^\omega A_\mathrm{p}(\tau')+p^\omega S_l^\omega(\tau',t)$$

which yields for any $\tau' \in [0,t]$ that

$$\mathsf{E}\left[D_{\mathrm{p}}(t)\right] \leq A_{\mathrm{p}}(\tau') + \sum_{\omega\in\Omega} p^{\omega} S_l^{\omega}(\tau', t)$$

since $\sum_{\omega\in\Omega} p^{\omega} = 1$. It then follows that

$$\mathsf{E}\left[D_{\mathrm{p}}(t)\right] \leq \inf_{\tau'\in[0,t]} \left\{A_{\mathrm{p}}(\tau') + \mathsf{E}\left[S_l^{\omega}(\tau', t)\right]\right\}$$

which completes the proof. ∎

2) Tandem Systems: The end-to-end service process of a sequence of systems is given by the service processes of the individual systems $S^h(\tau,t)$ $(h = 1 \ldots H)$ as [8]

$$S^{\mathrm{net}}(\tau,t) = S^1 \otimes \cdots \otimes S^H(\tau,t). \tag{11}$$

Note the difference between the min-plus convolution in (11) and the definition of end-to-end available bandwidth as a simple minimum in (2). Assuming work-conserving systems, we substitute $S^h(\tau,t)/(t-\tau) = \alpha^h(\tau,t)$ from (10). For the special case of constant-rate functions $S^h(\tau,t)/(t-\tau) = r^h$, it is shown in [28] that the min-plus convolution simplifies to $S^{\mathrm{net}}(\tau,t)/(t-\tau) = \min_h\{r^h\}$, which coincides with the available bandwidth $\alpha^{\mathrm{net}}(\tau,t) = \min_h\{r^h\}$. Without constant-rate functions, we only get $S^{\mathrm{net}}(\tau,t)/(t-\tau) \leq \alpha^{\mathrm{net}}(\tau,t)$. Here, we prove that equality can be recovered if the observed interval tends to infinity. We assume that the hth system has a long-term average service rate α_∞^h, i.e., for any τ, we have

$$\liminf_{t\to\infty} \frac{S^h(\tau,t)}{t-\tau} = \alpha_\infty^h. \tag{12}$$

Lemma 4: Given the end-to-end service process $S^{\mathrm{net}}(\tau,t)$ as in (11) where the service processes of the individual systems $S^h(\tau,t)$ satisfy (12). For any τ, it holds that

$$\liminf_{t\to\infty} \frac{S^{\mathrm{net}}(\tau,t)}{t-\tau} = \min_h \left\{\alpha_\infty^h\right\}.$$

Proof: It is sufficient to show the proof for two systems. Due to the properties of the convolution, the result applies to an arbitrary number of systems by repeated application. We rewrite (11) as

$$\frac{S^{\mathrm{net}}(\tau,t)}{t-\tau} = \inf_{\theta\in[\tau,t]} \left\{\frac{S^1(\tau,\theta)}{t-\tau} + \frac{S^2(\theta,t)}{t-\tau}\right\}.$$

Now we let $t\to\infty$. If $\alpha_\infty^1 > \alpha_\infty^2$, the minimum will be attained for finite θ such that the first term of the sum goes to zero and the second term to α_∞^2. Otherwise, if $\alpha_\infty^1 < \alpha_\infty^2$, the parameter θ will tend to t such that the first term becomes α_∞^1, and the second term goes to zero. This gives us

$$\liminf_{t\to\infty} \frac{S^{\mathrm{net}}(\tau,t)}{t-\tau} = \min\left\{\alpha_\infty^1, \alpha_\infty^2\right\}.$$

Finally, if $\alpha_\infty^1 = \alpha_\infty^2$, the result holds trivially. ∎

D. Max-Plus Network Calculus for Timestamps

So far, we have used the min-plus formulation of the network calculus, where arrival and departure functions $A(t)$ and $D(t)$ denote amounts of traffic in a time interval $(0,t]$. This formulation allowed us to establish the link between available bandwidth and service curves. There is an alternate (equivalent) formulation of the network calculus based on a max-plus algebra, which describes arrivals and departure functions in terms of timestamps [2], [8]. Since any probing scheme is based on taking timestamps of probing traffic, adopting a max-plus algebra for the computation of the available service is more convenient. Let $T_{\mathrm{A}}(n)$ and $T_{\mathrm{D}}(n)$ denote the timestamps of the nth probe packet. For unit-sized packets, functions $A(t)$ and $T_{\mathrm{A}}(n)$ are related by $A(t) = \sum_{n=1}^{\infty} 1_{\{T_{\mathrm{A}}(n)\leq t\}}$, where the indicator function $1_{\{X\leq x\}} = 1$ if $X \leq x$, and zero otherwise.[2]

We next show how an ε-effective service curve can be estimated from packet timestamps by switching to the max-plus representation of the network calculus. An additional benefit of using the max-plus representation is that it can deal with packet loss more easily by viewing lost packets as being infinitely delayed. We next show that min-plus and max-plus representations yield identical results.

Similar to (6), an ε-effective service curve that operates directly on packet timestamps can be defined in the max-plus algebra as

$$\mathsf{P}\left[T_{\mathrm{D}}(n) \leq \max_{\nu\in[1,n]} \left\{T_{\mathrm{A}}(\nu) + T_{\mathcal{S}}^{\varepsilon}(n-\nu)\right\}\right] \geq 1-\varepsilon. \tag{13}$$

Here, $T_{\mathcal{S}}^{\varepsilon}(n-\nu)$ specifies a shift-invariant upper bound on the amount of time required to serve $n-\nu+1$ packets. Under similar assumptions as in Theorem 1, a max-plus ε-effective service curve can be estimated from the delay defined as $W(n) = T_{\mathrm{D}}(n) - T_{\mathrm{A}}(n)$. Again, we use constant-rate arrivals $T_{\mathrm{A}}(n) = n/r$, where $r \in R$, and let $n\to\infty$ to observe the steady-state delay quantile $W^{\xi}(r)$. It can be shown that for all $n \geq 0$

$$T_{\mathcal{S}}^{\varepsilon}(n) = \min_{r\in R}\left\{\frac{n}{r} + W^{\xi}(r)\right\} \tag{14}$$

is an ε-effective service curve that satisfies (13) with violation probability $\varepsilon = \sum_r \xi$. We provide the proof in [32] and omit it here as it closely follows the proof of Theorem 1. Similar to Theorem 1, (14) lays the foundation for an estimation method that, however, takes delay measurements to estimate a max-plus service curve. The relation of the two methods is established by the following theorem, which states that the max-plus service curve estimate from (14) deviates from the min-plus service curve estimate from Theorem 1 by at most one packet length. The theorem uses the pseudo-inverse of service curve $T_{\mathcal{S}}$, which is defined as

$$\left(T_{\mathcal{S}}^{\varepsilon}\right)^{-1}(t) = \sum_{n=0}^{\infty} 1_{\left\{T_{\mathcal{S}}^{\varepsilon}(n)\leq t\right\}}.$$

Theorem 2: Given a system with ε-effective service curves $\mathcal{S}^{\varepsilon}(t)$ from Theorem 1 and $T_{\mathcal{S}}^{\varepsilon}(n)$ from (14), and assuming that the system forwards traffic in order of its arrival, the following holds:

$$\left(T_{\mathcal{S}}^{\varepsilon}\right)^{-1}(t) - 1 \leq \mathcal{S}^{\varepsilon}(t) \leq \left(T_{\mathcal{S}}^{\varepsilon}\right)^{-1}(t).$$

The proof of Theorem 2 uses the following fundamental lemma that relates the backlog of a system to its delay.

[2]Considering variable-sized packets requires additional notation to specify packet lengths. Since many probing methods use probe packets of fixed size, the assumption of unit sized packets is justified.

Lemma 5: Given a system with strictly increasing arrival and departure timestamps $T_{\mathrm{A}}(n)$ and $T_{\mathrm{D}}(n)$, respectively. Assume the system serves arrivals in order. The backlog observed at the departure time $T_{\mathrm{D}}(n)$ of packet n equals

$$\begin{aligned} B\left(T_{\mathrm{D}}(n)\right) &= A\left(T_{\mathrm{D}}(n) - W(n), T_{\mathrm{D}}(n)\right) \\ &= A\left(T_{\mathrm{A}}(n), T_{\mathrm{A}}(n) + W(n)\right) \end{aligned}$$

where $A(\tau, t)$ are the cumulative arrivals in $(\tau, t]$ and $W(n)$ is the delay of packet n.

Proof: From the definition of backlog, we have

$$B\left(T_{\mathrm{D}}(n)\right) = A\left(T_{\mathrm{D}}(n)\right) - D\left(T_{\mathrm{D}}(n)\right).$$

Since the arrivals are served in order, it holds that $D(T_{\mathrm{D}}(n)) = A(T_{\mathrm{A}}(n))$, and it follows by substitution that

$$B\left(T_{\mathrm{D}}(n)\right) = A\left(T_{\mathrm{A}}(n), T_{\mathrm{D}}(n)\right).$$

Using the definition of delay, $W(n) = T_{\mathrm{D}}(n) - T_{\mathrm{A}}(n)$ completes the proof. ∎

We use Lemma 5 to relate quantiles of the backlog and delay to each other, e.g., $B^{\xi} = rW^{\xi}$ for constant-rate arrivals with rate r, where we denote by B and W the steady-state backlog and delay for $n \to \infty$. We note that Little's law, i.e., $\mathsf{E}[B] = \lambda \mathsf{E}[W]$ for arrivals with average rate λ, can be recovered from Lemma 5; see the technical report [31]. Equipped with Lemma 5, we now prove Theorem 2.

Proof (of Theorem 2): Consider the argument of the indicator function $T_{\mathcal{S}}^{\varepsilon}(n) \leq t$ in the definition of the pseudo-inverse $\left(T_{\mathcal{S}}^{\varepsilon}\right)^{-1}$. By insertion of $T_{\mathcal{S}}^{\varepsilon}(n)$ from (14), we have $\min_{r \in R}\{n/r + W^{\xi}(r)\} \leq t$ and after some reordering

$$n \leq \max_{r \in R} \left\{rt - rW^{\xi}(r)\right\}.$$

Instantiating Lemma 5 with $A(t) = \lfloor rt \rfloor$ yields the backlog $B(T_{\mathrm{D}}(n)) = \lfloor rW(r,n) \rfloor$. Letting $n \to \infty$ and taking quantiles, we obtain $B^{\xi}(r) = \lfloor rW^{\xi}(r) \rfloor = rW^{\xi}(r) - \vartheta(r)$, where $\vartheta(r) \in [0, 1)$. It follows that the condition $T_{\mathcal{S}}^{\varepsilon}(n) \leq t$ is equivalent to

$$n \leq \max_{r \in R} \left\{rt - B^{\xi}(r) - \vartheta(r)\right\}.$$

With Theorem 1 and since $n = 0, 1, 2, \ldots$ is an integer, the claim follows. ∎

IV. Derivation of the Probing Methodology

We next address how the results from Section III can be developed into practical probing schemes for networks with random service. The probing schemes are expressed as described in Section III-D, where arrival and departure timestamps of packet train probes are used to measure steady-state delay quantiles and to compute an estimate of the service curve.

We describe a measurement experiment with constant-rate packet train probes by a tuple $\langle R, N, I \rangle$, where R is the set of probing rates, N is the number of packets in a train, and I is the number of repeated measurements. The selection of these parameters is crucial for a probing scheme, yet the derivations in Section III do not provide guidelines for selecting values. For example, a single reading of the steady-state delay $W(r)$ from a packet train is ideally taken from an infinitely long packet train ($N \to \infty$). Also, computing the exact tail distribution of the delay requires an infinite number of repeated measurements ($I \to \infty$). Clearly, such ideal measurement experiments are not viable. In this section, we will develop and evaluate guidelines for practical choices for $\langle R, N, I \rangle$. In Section IV-A, we discuss the selection of the rates. In Section IV-B, we show how to obtain stationary delay estimates using finite packet trains and a finite number of repeated measurements with the usage of statistical methods. In Section IV-C, we discuss how to extract information from short packet trains. In Section IV-D, we apply our method to nonlinear and lossy networks.

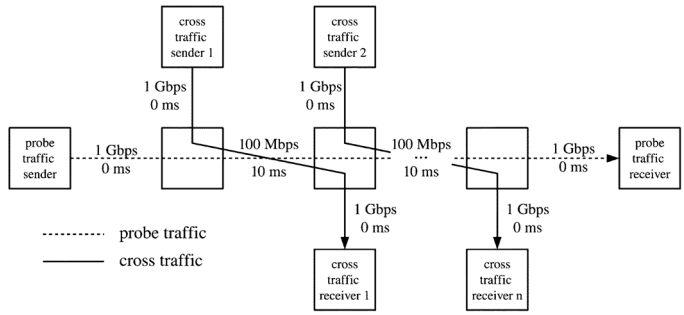

Fig. 2. Multihop network with multiple 100-Mbps bottleneck links.

For an experimental evaluation, we resort to a local testbed as well as the Emulab testbed [43], which offer controlled experiments on real networking equipment. We consider the topology shown in Fig. 2, where probe traffic and cross traffic are multiplexed at a tandem of 100-Mbps bottleneck links. The capacities and delays in the network are specified in the figure. We consider different packet schedulers, including priority, fair queueing, and FIFO, and different buffer sizes at the bottleneck links. The default configuration is a network with a single bottleneck link and priority scheduling with high priority to cross traffic and a large buffer size (of 10^6 packets). Cross traffic has a mean rate of 50 Mbps and consists of equally spaced packet bursts of back-to-back packets whose size follows a truncated Exponential or Pareto distribution. A comparison of outcomes for these distributions reveals the sensitivity of the measurement methodology to the burstiness of network traffic. The average size of a packet burst is 1500 B, and the shape parameter of the Pareto distribution is 1.5. With the chosen link capacity and cross-traffic rate, the limiting rate of the service process (= long-term available bandwidth) is 50 Mbps. We use D-ITG [4] to generate cross traffic. D-ITG generates truncated arrival distributions to conform to the maximum IP payload size of 64 kB. The payload is further divided by IP fragmentation to packet sizes of at most 1500 B. For probe traffic, we use Rude/Crude [25], which emits constant-bit-rate packet trains consisting of packets with a size of 1500 B. NTP is used for time synchronization. We note that clock synchronization over production networks is challenging [40]. While a constant clock offset mainly results in a time shift of the service curve estimate [28], drifting clocks may cause distortions.

A. Selection of Probing Rates

The selection of the set of probing rates presents a tradeoff. On the one hand, adding probing rates improves the estimate of the ε-effective service curves since each rate contributes an additional linear segment (see Fig. 1). On the other hand, since we

compute the violation probability by an application of the union bound, adding probing rates increases the violation probability. We use an algorithm that seeks to find a small set of characteristic rates that contribute to the service curve. The algorithm combines a binary increase and a binary search algorithm, similar to the rate selection procedure in [17].

The algorithm has as parameter r_{acc} to specify the desired rate resolution. The binary increase starts at $r_1 = r_{\mathrm{acc}}$. As long as the probes at rate r_i measure a finite delay quantile, the rate is doubled. The estimation of the delay quantile is explained in Section IV-B. The first rate r_i, at which no finite delay quantile can be measured, is used to start a binary search in the interval $[r_{i-1}, r_i]$ using the same test criterion. Each additional probing rate halves the interval. Once the interval achieves the target accuracy r_{acc}, the rate scan is terminated. Let $\tilde{r}$ be the largest rate that achieves a finite delay quantile. As an example, if $\tilde{r} = 50$ Mbps and $r_{\mathrm{acc}} = 4$ Mbps, the algorithm probes at rates $r_i = 4, 8, 16, 32, 64, 48, 56$, and 52 Mbps and stops at the interval [48, 52].

The total number of rates probed by the binary increase/binary search algorithm is $2\lfloor \log_2(\tilde{r}/r_{\mathrm{acc}}) \rfloor + 2$, where the binary increase algorithm requires $\lfloor \log_2(\tilde{r}/r_{\mathrm{acc}}) \rfloor + 2$ steps until it first exceeds $\tilde{r}$, and the binary search performs another $\lfloor \log_2(\tilde{r}/r_{\mathrm{acc}}) \rfloor$ steps to ensure the target accuracy.

B. Estimation of Steady-State Delays

Next, we discuss how to estimate the tail distribution of steady-state delays using finite-length packet train probes. We employ a statistical test to detect stationarity of a time series, and use it to adapt the length of packet train probes to the variability of cross traffic.

1) Stationarity Test: By definition, (14) uses steady-state delays to compute a service curve. While Lemma 2 states that the steady-state delay distribution exists as long as the rate of probe arrivals does not exceed the limiting rate of the service process, reaching the steady state requires infinitely long packet trains. To determine the steady-state delay from a finite-length packet train, we use a statistical test that detects if the delays $W(n)$ of a sequence of probe packets $n \in [0, N-1]$ satisfy stationarity. If stationarity is detected, we use the delay of the last packet of a packet train as an estimate of the steady-state delay. The delay values of all other packets from the same packet train are discarded due to possible correlations. If stationarity cannot be detected, we set the delay estimate to infinity. We repeat the measurement I times for each probing rate r to measure the $(1-\xi)$-quantile of the delay $W^{\xi}(r)$. Note that due to the minimum in (14), delay quantiles of infinity do not contribute to the service curve estimate.

To detect stationarity of the delay series observed by a packet train, we use the unit root test from Elliot, Rothenberg, and Stock (ERS) [12]. The *ERS test* is based on an auto-regressive moving average model. The null hypothesis of the test is that the data series has a unit root, which implies nonstationarity. The ERS test returns a single value referred to as *ERS statistic*. If the ERS statistic falls below a critical value, the null hypothesis is rejected, and stationarity is assumed.

2) Adaptive Train Length: Since the minimal train length that permits detecting stationarity is not known *a priori*, we define a procedure that adaptively increases the train length. When the ERS test indicates nonstationarity for a share of more than ξ of the packet trains sent at a certain rate, then either the stationary $(1-\xi)$-delay quantiles cannot be achieved at this probing rate, or the train length is too short to reliably detect stationarity. To assess whether increasing the train length can help, we inspect the trend of the ERS statistic. We compute the ERS statistic for the first half of the train and for the entire train. If the ERS statistic decreases, i.e., if the ERS statistic of the first half of the train is larger than the ERS statistic of the entire train, we interpret this as an indication that longer trains may achieve stationarity. We refer to this test as the *trend test* that is passed by a packet train if its ERS statistic decreases. If the majority of the packet trains sent at a certain rate passes the trend test, we double the train length and carry out the measurements at this rate anew. We repeat this procedure until either stationarity is achieved or the majority of the trains fails the trend test.

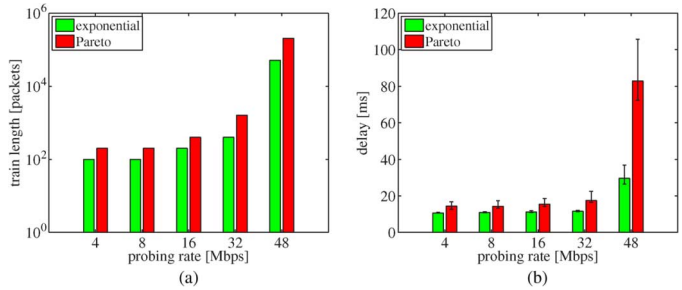

Fig. 3. (a) Required train length needed to observe stationary delays at a link with Exponential and Pareto cross traffic and a limiting rate of 50 Mbps. For the train lengths in (a), (b) presents the observed delay values as stationary 0.95-delay quantiles (with 0.95 confidence intervals).

We next show the train length required to achieve stationarity at a certain probing rate. Fig. 3(a) shows the train length that permits detecting stationarity for a share of at least $(1-\xi) = 0.95$ of $I = 250$ packet trains sent at different probing rates and for different types of cross traffic each. The probing rates are chosen according to the algorithm from Section IV-A, and the train length is adapted as described above starting at a minimum train length of 100 packets. The results show that the required train length is sensitive to the distribution of cross traffic and to the probing rate. The required train length increases sharply, when the probing rate approaches the limiting rate of 50 Mbps.

3) Tail Distribution: The computation of ε-effective service curves requires the $(1-\xi)$-quantiles of the stationary delays, $\widetilde{W}^{\xi}(r)$. Since each packet train only provides one reading of the stationary delays, obtaining a delay quantile requires multiple packet trains for each probing rate. In particular, to compute a 0.95-delay quantile, the minimum number of repeated measurements for a given rate is $I = 20$.

The delays observed by different packet trains can be assumed to be independent when packet trains have random start times (see [3] for a discussion). We can quantify the accuracy of the delay quantiles using confidence intervals (which, for quantiles, are computed from the binomial distribution). This allows us to increase the number of packet trains until a desired accuracy is met. Fig. 3(b) displays the stationary 0.95-delay quantiles and corresponding 0.95 confidence intervals achieved with $I = 250$ repeated measurements [where packet train lengths are given in Fig. 3(a)]. The indicated confidence intervals show that the accuracy decreases when the probing rate approaches the limiting rate and when cross traffic is more bursty. To reduce the

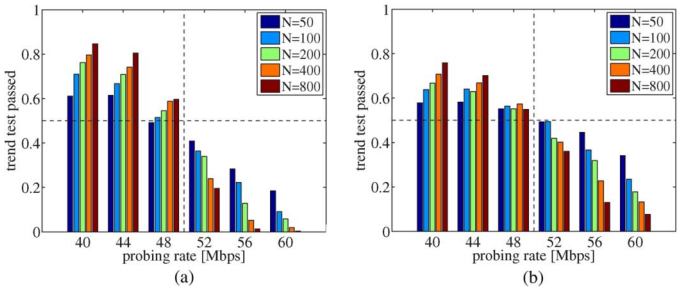

Fig. 4. Fraction of packet trains, out of $I = 1000$ trains, that pass the trend test for different probing rates and given packet train length $N = 50$–800. Packet trains with a rate below the limiting rate of 50 Mbps are classified correctly if they pass the trend test, and packet trains with a higher rate are classified correctly if they fail the trend test. (a) Exponential cross traffic. (b) Pareto cross traffic.

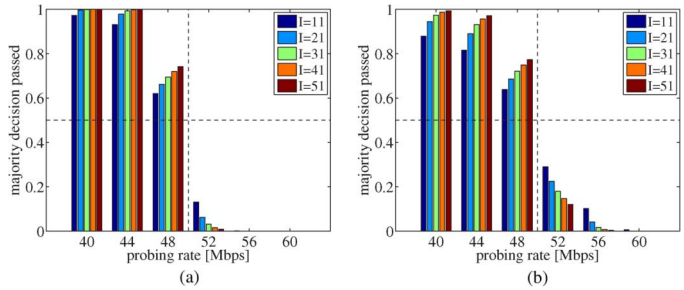

Fig. 5. Outcome of a majority decision on the trend test using I repeated measurements for a train length of $N = 200$ packets. As in Fig. 4, trains with a rate below 50 Mbps should pass the trend test, and packet trains with a rate above 50 Mbps should fail the trend test. (a) Exponential cross traffic. (b) Pareto cross traffic.

number of repeated measurements, we refer to the technical report [31] for methods to predict the tail of the delay distribution.

C. Short Packet Train Probes

As seen in Fig. 3(a), when the network is close to saturated, measuring stationary delays may require very long packet trains. In this section, we discuss how the limiting rate and a service curve can be estimated with small sized packet trains, without requiring stationary delays.

1) Limiting Rate Estimation: In principle, the limiting rate of a service curve estimate is the largest probing rate that observes finite steady-state delays. In practice, however, the required packet train length needed to read steady-state delays may become prohibitive. Thus, we abandon trying to read steady-state delays in this regime and, instead, estimate the limiting rate directly. Based on the ERS test and the trend test, we devise a heuristic method to obtain the limiting rate with packet trains with at most N packets.

In Section IV-B.2, we used the ERS test to detect stationarity of delays, and the trend test to indicate that stationarity may be observable for longer packet trains (sent at the same rate). To limit the train length at N, we now consider passing the trend test as sufficient to assume that a finite steady-state delay exists at a given rate. When the trend test fails, we assume that no finite steady-state delay exists. We use the interpretation of the trend test, i.e., finite steady-state delays exist or do not exist, as criteria in the binary increase/binary search algorithm from Section IV-A for increasing or decreasing the probing rate. Then, the largest probing rate for which the trend test passes is an estimate of the limiting rate.

In Fig. 4, we evaluate the impact of the train length on the fidelity of the trend test. We use again the network shown in Fig. 2. Since the network has a limiting rate of 50 Mbps, packet trains should pass the trend test if the probing rate is below 50 Mbps, and fail the trend test if the probing rate exceeds 50 Mbps. The bars in Fig. 4 depict the fraction of trains that pass the trend test out of a total of $I = 1000$ trains, for probing rates $r = 40$–60 Mbps. We show results for packet trains with a length of $N = 50$–800. The figure shows that the likelihood of a correct classification increases when the train length is increased, and that it decreases when the probing rate approaches the limiting rate and when the cross traffic is more bursty. On the other hand, note that the majority of trains make a correct classification.

We can exploit the last observation to achieve a robust classification when the number of packet trains is small. We perform measurements of I packet trains, perform the trend test for each train, and then conduct a majority decision. By selecting I to be an odd number, we can ensure that that a majority decision is always feasible. Assuming independence of packet trains, we compute the probability that a majority of trains passes the test from Fig. 4 via the binomial distribution. Fig. 5 shows the results for a train length of $N = 200$ packets and $I = 11$–51 repeated measurements. Even for Pareto traffic, using a majority decision of the trend tests results in a correct classification.

2) Service Curve Estimation: Short packet trains often do not pass the ERS test and, therefore, do not allow a reading of stationary delays. It is possible to construct a service curve with nonstationary delays, which, however, is only valid for time intervals with limited range. Using the procedure from Section IV-C, we can measure the delay of the last packet (with index N) of each train and compute a service curve estimate $T^{\varepsilon}_{\mathcal{S}}(n)$ with (14) for the range $n \in [0, N-1]$. The details are described in the technical report [31]. By selecting N sufficiently large, the range of validity of the service curve can be made large enough to cover the timescales of interest.

For the same network as before, Fig. 6 depicts service curves obtained with packet trains of length $N = 800$–$12\,800$ and compares them to a service curve estimate with unrestricted train lengths (which observes stationary delays). The probing rates are set for a target resolution of $r_{\mathrm{acc}} = 4$ Mbps, and the number of repeated measurements is $I = 250$. Fig. 6 shows the computed service curves. In the figure, the limiting rate of 50 Mbps is indicated by the dashed diagonal lines. It is evident that all service curve estimates closely track the limiting rate of 50 Mbps. For Exponential cross traffic, service curves computed with $N = 800$ provide similar results as unrestricted packet trains. For the burstier Pareto traffic, we observe that the service curve segments computed by short trains sometimes overestimate the service curve estimate obtained from stationary delays. Here, when the probing rate approaches the limiting rate, the delay estimates obtained from short trains underestimate the stationary delay distribution, which results in an overestimation of the service curve.

D. Nonlinear and Lossy Networks and Elastic Cross Traffic

So far, we assumed that network elements such as links, queues, and schedulers can be modeled by (5) as lossless

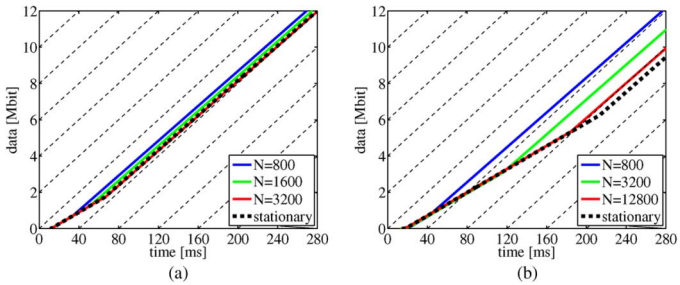

Fig. 6. Service curve estimates from trains of length N compared to unrestricted trains that observe stationary delays. (a) Exponential cross traffic. (b) Pareto cross traffic.

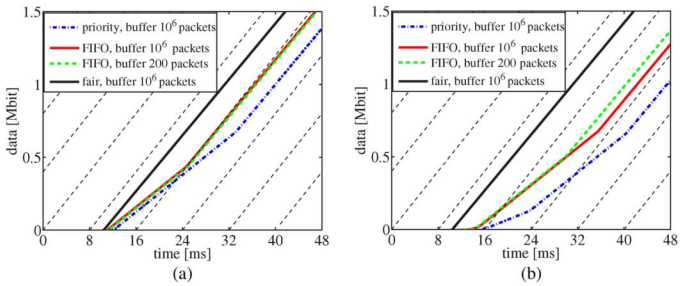

Fig. 7. Service curve estimates for priority, fair, and FIFO scheduling with large and small buffer. (a) Exponential cross traffic. (b) Pareto cross traffic.

linear systems. Next, we show results for systems where these assumptions are relaxed.

A FIFO scheduler is a prime example of a nonlinear system. As shown in [13], a FIFO scheduler operates as a linear system only when the traffic (from probe and cross traffic) is less than the capacity of the system. When it becomes overloaded, the FIFO scheduler enters into a nonlinear regime. Thus, the application of linear system theory to FIFO schedulers is justified only when the capacity is not saturated. In our case, since the stationarity test ensures that the probing traffic does, on average, not exhaust the capacity of the network, the linearity assumption is often justified. Bursty arrivals, however, cause short-term violations of linearity.

The assumption of lossless systems can be easily relaxed in a max-plus algebra by simply modeling a dropped packet as incurring an infinite delay. That is, if packet n has been dropped, we set $T_{\mathrm{D}}(n) = \infty$. As a result, probing rates experiencing a packet-loss ratio of ξ or more do not contribute to the service curve estimate.

In Fig. 7, we show service curve estimates obtained for a link with fair and priority scheduling, both of which are linear, to FIFO scheduling, which is nonlinear in overload. For priority scheduling, we assume that cross traffic has a higher priority. Buffers are large (10^6 packets) enough so that no packet losses occur. For a FIFO scheduler, we also include results with a small buffer size of 200 packets, which results in moderate packet losses (the loss is less than 1% at the limiting rate). The probing parameters are $r_{\mathrm{acc}} = 4$ Mbps, $I = 250$, $\xi = 0.05$, and $N = 800$. For all scenarios, the service curves exhibit a latency of slightly above 10 ms, which matches the propagation delay of the bottleneck link (see Fig. 2). The service curve estimate for the fair scheduler is a straight line since the scheduler allocates a fair share of 50 Mbps to the probe traffic regardless of the burstiness of cross traffic. The service curve estimate at a priority scheduler, on the other hand, is sensitive to the type of cross traffic. Here, the higher-priority cross traffic results in additional latencies and a lower initial rate of the service curve. For the FIFO scheduler, we observe that the service curve estimates with small buffers and large buffers are very close, indicating that our estimation method deals well with packet losses. For Pareto cross traffic and with small buffers, the additional packet losses compared to Exponential cross traffic result in a larger estimate of the service curve.

TABLE I
LIMITING RATE ESTIMATES WITH ELASTIC CROSS TRAFFIC

cross traffic flows	exponential		Pareto	
	limiting rate [Mbps]	latency [ms]	limiting rate [Mbps]	latency [ms]
1 TCP	48	15	48	22
10 TCP	48	16	48	23
100 TCP	48	16	48	24
1 UDP	48	19	48	26

We also evaluated the estimation method in a network where cross traffic consists of elastic TCP traffic using the TCP Cubic algorithm. Since TCP reacts to congestion in the network, probing traffic could displace TCP traffic and overestimate the available bandwidth. We compare service curve estimates made with 1 UDP cross-traffic flow to estimates with 1, 10, and 100 TCP cross-traffic flows. In each case, the average rate of cross traffic is 50 Mbps, using the same parameters as before ($r_{\mathrm{acc}} = 4$ Mbps, $I = 250$, $\xi = 0.05$, $N = 800$). The scheduling is FIFO with a buffer size of 10^6 packets. Traffic generation at the application follows an Exponential or Pareto distribution. Table I depicts the initial latency and the limiting rate of the computed service curve (this corresponds to expressing the service curve to a simple latency-rate function). For all cases, the same limiting rate of 48 Mbps is computed. Also, the estimates of the latencies depend only on the traffic distribution (Pareto or Exponential), but not on the composition of the traffic. This indicates that TCP traffic is not displaced by the estimation method.

V. COMPARATIVE EVALUATION

Here, we compare the stochastic bandwidth estimation method with bandwidth estimation methods from the literature. We consider the network topology in Fig. 2, with FIFO scheduling, and parameters as discussed in Section IV. We also present measurements of an IEEE 802.11a network.

A. Service Curve Estimation

We first perform a comparison to a probing method for min-plus time-invariant deterministic linear systems from [28] for the network in Fig. 2 with a single bottleneck with FIFO scheduling and a buffer size of 10^6 packets. The method uses constant-rate packet trains with fixed length ($N = 800$ packets), where the rate of the packet trains is incremented between successive trains by 8 Mbps. We match these parameters for the stochastic bandwidth estimation by using adaptively varied train length with up to 800 packets and a target accuracy of $r_{\mathrm{acc}} = 8$ Mbps. Also, we set $\xi = 0.05$. For both methods, we present results from $I = 200$ iterations. The stochastic bandwidth estimation uses repeated measurements to obtain $(1 - \xi)$-delay quantiles, as well as 0.95 confidence intervals of the delays to derive a

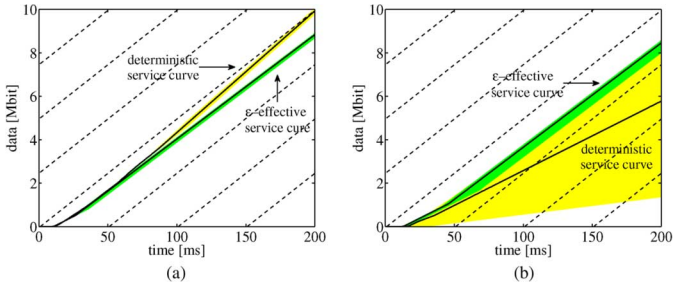

Fig. 8. Comparison of stochastic bandwidth estimation and deterministic bandwidth estimation. The ε-effective service curve is compared to the average deterministic service curve, with 0.95 confidence intervals indicated by shaded areas. (a) Exponential cross traffic. (b) Pareto cross traffic.

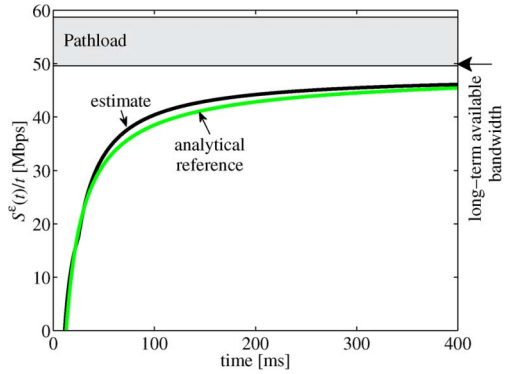

Fig. 9. Comparison of estimated ε-effective service curve with analytical reference of ε-effective service curve and estimates by Pathload tool (service curves are displayed as rate functions to allow for comparison).

single estimate of an ε-effective service curve. Since the method for deterministic time-invariant systems [28] generates one deterministic service curve estimate in each iteration, we compute the mean and 0.95 confidence intervals for all iterations.

Fig. 8 shows the average value of the deterministic service curve and the ε-effective service curve. Confidence intervals are shown as shaded areas. With Exponential cross traffic, both deterministic and stochastic service curve estimates give comparable results. A closer comparison to the diagonal reference lines shows that the deterministic service curve overestimates the limiting rate of 50 Mbps. For Pareto cross traffic, the deterministic service curve results in a lower estimate and very large confidence intervals. The ε-effective service curve recovers the limiting rate closely with small confidence intervals. Note that the service curves capture the non-work-conserving aspect of the available service in terms of the initial latency.

B. Bandwidth Estimation Tools

Most existing bandwidth estimation methods seek to find the long-term available bandwidth as defined in Section II. We present an example to illustrate that the long-term available bandwidth only presents part of the information on the available service in a network.

We consider a single 100-Mbps FIFO bottleneck link with Exponential cross traffic sent at rate 50 Mbps. We set the buffer limit to 200 packets (experiments with larger buffers yielded similar results). The other network and probing parameters are as in Section IV-D.

For the example, we select the Pathload [17] tool, which is frequently used as a benchmark method. The method reports a lower and an upper bound of the available bandwidth. (We refer to our technical report for a comparison with IGI/PTR [16], Spruce [42], and dietTOPP [22].) As a reference, we also include an analytically computed ε-effective service curve, which is computed as a leftover service curve [7] using a sample path bound for the cross traffic. For Exponential cross traffic, the analytical curve is computed with the Erlang distribution and the union bound.

Fig. 9 presents the results. Since the available bandwidth in (1) is defined as a rate, we convert the estimated and analytical ε-effective service curves $\mathcal{S}^{\varepsilon}(t)$ to a rate function $\mathcal{S}^{\varepsilon}(t)/t$. We note that the estimated ε-effective service curve matches the analytical reference very closely. The gray area shows the median of 100 repeated Pathload measurements. The figure shows that Pathload recovers the long-term available bandwidth well. On the other hand, the analytical and estimated ε-effective service curves express that significantly fewer resources are available on shorter timescales, which is due to the short-term randomness of cross traffic and link delays.

C. Multiple Bottleneck Links

We now address networks with multiple bottleneck links. In such networks, many estimation methods have been reported to severely underestimate the long-term available bandwidth on an end-to-end path [19]. In Section III-C, we showed that the end-to-end available bandwidth from (2) is generally greater than the end-to-end service process of a network given by (11). It recovers the long-term available bandwidth if the observation duration tends to infinity.

Since these results suggest that longer packet trains provide better estimates, we investigate how the length of packet trains impacts the accuracy of end-to-end available bandwidth estimates. We use the topology in Fig. 2, with one, three, and five 100-Mbps bottleneck links with FIFO scheduling and large buffers (of 10^6 packets). At each bottleneck link, we have independent Exponential cross traffic with an average rate of 50 Mbps. The remaining parameters are as defined at the end of Section V-B. We report results of the limiting rate for the ε-effective service curve estimate from Section IV-C with a target accuracy of $r_{\mathrm{acc}} = 1$ Mbps and the available bandwidth range reported by Pathload, where we show the median of 100 trials. We vary the maximum length of a packet train from 100 to 1600 packets. Since Pathload by default uses packet trains with a fixed length of 100 packets, we modified the source code of the tool.

Table II shows the limiting rates of the ε-effective service curve estimates and the available bandwidth bounds computed by Pathload. With short packet trains, estimates for a single bottleneck link are accurate. However, increasing the number of bottleneck links results in lower estimates. Longer packet trains provide better results for multiple bottleneck links, with both compared methods. This leaves open the possibility that underestimation of the long-term available bandwidth can be remedied by increasing the length of packet train probes.

D. IEEE 802.11a WiFi Networks

We apply the stochastic bandwidth estimation to a WiFi network, where the medium access control of the DCF results in a

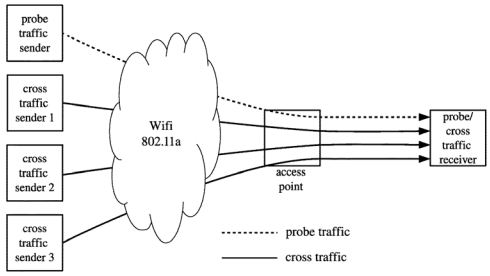

Fig. 10. Wireless network with several cross-traffic senders.

TABLE II
AVAILABLE BANDWIDTH ESTIMATES FOR MULTIPLE BOTTLENECK LINKS

train length [packets]	ε-eff. serv. curve (limiting rate) [Mbps]			Pathload (available bandwidth) [Mbps]					
				lower bound			upper bound		
	bottleneck links			bottleneck links			bottleneck links		
	1	3	5	1	3	5	1	3	5
100	48	43	41	50	43	38	58	53	44
200	48	45	44	51	45	41	54	52	45
400	48	47	46	51	47	43	53	51	45
800	48	47	47	51	49	45	52	51	46
1600	48	48	48	51	49	46	51	51	47

nondeterministic and non-work-conserving channel allocation.[3] We measure an IEEE 802.11a testbed network with and without cross traffic as shown in Fig. 10. Probe and cross traffic compete for the wireless medium using the DCF. All transmitted packets are received by an access point, which forwards the packet to a receiver station on a wireline 100-Mbps link. Since the DCF has been previously related to a fair queueing system [6], we compare the results to a wired network that employs fair queueing at a bottleneck link. The capacity of the wired bottleneck link is throttled to 30 Mbps. This corresponds to the maximum achievable throughput in the WiFi 802.11a network [5]. As we seek to present results on the DCF, we avoid additional randomness by using constant-rate cross traffic with a rate of 25 Mbps. The rate is divided equally among the cross traffic flows, so that each flow uses at least its fair share. This allows exploring the behavior of the DCF in a fully utilized IEEE 802.11a network. The parameters for the packet trains are $N = 800$ packets, $I = 250$ iterations, $\xi = 0.05$, and a target accuracy of $r_{\text{acc}} = 1$ Mbps. Fig. 11 shows the rate of the service curve estimates $\mathcal{S}^{\varepsilon}(t)/t$ for the WiFi network and the wireline fair queueing link. For the fair scheduler, without cross traffic, the rate quickly approaches the bound of 30 Mbps. When cross traffic is added, the rates track the fair share of 15 and 7.5 Mbps for one and three cross-traffic flows, respectively. For the WiFi network, the limiting rates estimates do not quite reach the computed fair share. Also, the slower convergence to the limiting rate reflects additional delays in the WiFi network. The lower rate and higher delays seen in the WiFi network are due to the accumulated impacts of the random backoff of the DCF, collisions, and retransmissions.

VI. CONCLUSION

We have presented a system-theoretic foundation for bandwidth estimation of networks with random service, where we used the framework of the stochastic network calculus to derive a method that estimates an ε-effective service curve from steady-state backlog or delay quantiles observed by packet train probes. The service curve model extends to networks of nodes, as well as to non-work-conserving systems. The ε-effective service curve characterizes service availability at different timescales and recovers the long-term available bandwidth as its limiting rate. While ideal measurements require an infinite repetition of infinitely long packet trains, we showed that practical estimation methods can be achieved by applying statistical tests and using appropriate heuristics. We found that cross-traffic variability, the number of bottleneck links, and the target accuracy of the estimate have a significant impact on the amount of probes required. We presented measurement examples, which showed that estimates of ε-effective service curves can disclose essential characteristics of random service in wired and wireless networks, which are not reflected in the long-term available bandwidth.

[3]We note that [41] recently applied our method to compute service curves of a simulated wireless channel.

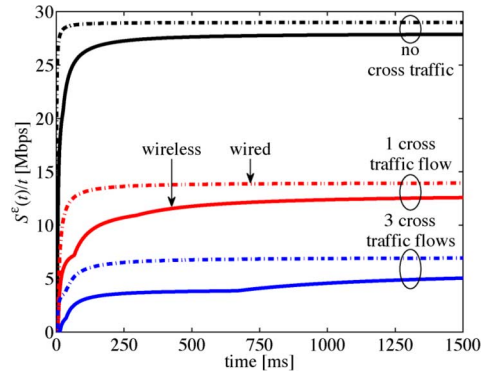

Fig. 11. Service curves as rate functions for a wireless IEEE 802.11a network, and a wireline 30 Mbps link with fair queueing. The graphs show the rate of the service curve estimates with 0, 1, and 3 cross traffic flows.

REFERENCES

[1] F. Agharebparast and V. C. M. Leung, "Slope domain modeling and analysis of data communication networks: A network calculus complement," in *Proc. IEEE ICC*, Jun. 2006, pp. 591–596.
[2] F. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat, *Synchronization and Linearity: An Algebra for Discrete Event Systems*. New York, NY, USA: Wiley, 1992.
[3] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot, "The role of PASTA in network measurement," in *Proc. ACM SIGCOMM*, Sep. 2006, pp. 231–242.
[4] A. Botta, A. Dainotti, and A. Pescape, "Multi-protocol and multi-platform traffic generation and measurement," presented at the IEEE INFOCOM, May 2007.
[5] M. Bredel and M. Fidler, "A measurement study of bandwidth estimation in IEEE 802.11g wireless LANs using the DCF," in *Proc. IFIP Netw.*, May 2008, pp. 314–325.
[6] M. Bredel and M. Fidler, "Understanding fairness and its impact on quality of service in IEEE 802.11," in *Proc. IEEE INFOCOM*, Apr. 2009, pp. 1098–1106.
[7] A. Burchard, J. Liebeherr, and S. Patek, "A min-plus calculus for end-to-end statistical service guarantees," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4105–4114, Sep. 2006.
[8] C.-S. Chang, *Performance Guarantees in Communication Networks*. Berlin, Germany: Springer-Verlag, 2000.
[9] A. Delphinanto, T. Koonen, and F. den Hartog, "Real-time probing of available bandwidth in home networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 134–140, Jun. 2011.
[10] B. K. Dey, D. Manjunath, and S. Chakraborty, "Estimating network link characteristics using packet-pair dispersion: A discrete-time queueing theoretic analysis," *Comput. Netw.*, vol. 55, pp. 1052–1068, Apr. 2011.

[11] S. Ekelin, M. Nilsson, E. Hartikainen, A. Johnsson, J.-E. Mangs, B. Melander, and M. Bjorkman, "Real-time measurement of end-to-end available bandwidth using Kalman filtering," in *Proc. IEEE/IFIP NOMS*, Apr. 2006, pp. 73–84.
[12] G. Elliott, T. J. Rothenberg, and J. H. Stock, "Efficient tests for an autoregressive unit root," *Econometrica*, vol. 64, no. 4, pp. 813–836, July 1996.
[13] Y. Ghiassi-Farrokhfal and J. Liebeherr, "Output characterization of constant-bit-rate traffic in FIFO networks," *IEEE Commun. Lett.*, vol. 13, no. 8, pp. 618–620, Aug. 2009.
[14] P. Hága, K. Diriczi, G. Vattay, and I. Csabai, "Understanding packet pair separation beyond the fluid model: The key role of traffic granularity," in *Proc. IEEE INFOCOM*, Apr. 2006, pp. 2374–2386.
[15] T. Hisakado, K. Okumura, V. Vukadinovic, and L. Trajkovic, "Characterization of a simple communication network using Legendre transform," in *Proc. IEEE ISCAS*, May 2003, pp. 738–741.
[16] N. Hu and P. Steenkiste, "Evaluation and characterization of available bandwidth probing techniques," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 6, pp. 879–894, Aug. 2003.
[17] M. Jain and C. Dovrolis, "Pathload: A measurement tool for end-to-end available bandwidth," in *Proc. PAM*, Mar. 2002, pp. 14–25.
[18] M. Jain and C. Dovrolis, "End-to-end available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput," *IEEE/ACM Trans. Netw.*, vol. 11, no. 4, pp. 537–549, Aug. 2003.
[19] M. Jain and C. Dovrolis, "Ten fallacies and pitfalls on end-to-end available bandwidth estimation," in *Proc. ACM IMC*, Oct. 2004, pp. 272–277.
[20] Y. Jiang and Y. Liu, *Stochastic Network Calculus*. New York, NY, USA: Springer, 2008.
[21] A. Johnsson and M. Bjorkman, "On measuring available bandwidth in wireless networks," in *Proc. IEEE LCN*, Oct. 2008, pp. 861–868.
[22] A. Johnsson, B. Melander, and M. Björkman, "Bandwidth measurement in wireless networks," in *Proc. Med-Hoc-Net*, Jun. 2005, pp. 89–98.
[23] P. Kanuparthy, C. Dovrolis, K. Papagiannaki, S. Seshan, and P. Steenkiste, "Can user-level probing detect and diagnose common home-WLAN pathologies," *Comput. Commun. Rev.*, vol. 42, no. 1, pp. 7–15, Jan. 2012.
[24] M. A. Y. Khan and D. Veitch, "Speedo: Realistic achievable bandwidth in 802.11 through passive monitoring," in *Proc. LCN WNM*, 2008, pp. 892–899.
[25] J. Laine, S. Saaristo, and R. Prior, "RUDE & CRUDE," 2002 [Online]. Available: http://rude.sourceforge.net/
[26] J.-Y. Le Boudec and P. Thiran, *Network Calculus A Theory of Deterministic Queuing Systems for the Internet*. Berlin, Germany: Springer-Verlag, 2001.
[27] J. Li, M. Claypool, and R. R. Kinicki, "WBest: A bandwidth estimation tool for IEEE 802.11 wireless networks," in *Proc. IEEE LCN*, Oct. 2008, pp. 374–381.
[28] J. Liebeherr, M. Fidler, and S. Valaee, "A system theoretic approach to bandwidth estimation," *IEEE/ACM Trans. Netw.*, vol. 18, no. 4, pp. 1040–1053, Aug. 2010.
[29] X. Liu, K. Ravindran, and D. Loguinov, "A queuing-theoretic foundation of available bandwidth estimation: Single-hop analysis," *IEEE/ACM Trans. Netw.*, vol. 15, no. 4, pp. 918–931, Aug. 2007.
[30] X. Liu, K. Ravindran, and D. Loguinov, "A stochastic foundation of available bandwidth estimation: Multi-hop analysis," *IEEE/ACM Trans. Netw.*, vol. 16, no. 1, pp. 130–143, Apr. 2008.
[31] R. Lübben, M. Fidler, and J. Liebeherr, "A foundation for stochastic bandwidth estimation of networks with random service," Tech. Rep. arXiv:1008.0050v1 [cs.NI], Jul. 2010.
[32] R. Lübben, M. Fidler, and J. Liebeherr, "A foundation for stochastic bandwidth estimation of networks with random service," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 1817–1825.
[33] S. Machiraju, D. Veitch, F. Baccelli, and J. Bolot, "Adding definition to active probing," *Comput. Commun. Rev.*, vol. 37, no. 2, pp. 19–28, Apr. 2007.
[34] B. Melander, M. Björkman, and P. Gunningberg, "A new end-to-end probing and analysis method for estimating bandwidth bottlenecks," in *Proc. IEEE GLOBECOM*, Nov. 2000, pp. 415–420.
[35] B. Melander, M. Björkmann, and P. Gunningberg, "First-come-first-served packet dispersion and implications for TCP," in *Proc. IEEE GLOBECOM*, Nov. 2002, pp. 2170–2174.
[36] K.-J. Park, H. Lim, and C.-H. Choi, "Stochastic analysis of packet-pair probing for network bandwidth estimation," *Comput. Netw.*, vol. 50, no. 12, pp. 1901–1915, May 2006.
[37] K.-J. Park, H. Lim, J. C. Hou, and C.-H. Choi, "Feedback-assisted robust estimation of available bandwidth," *Comput. Netw.*, vol. 53, no. 7, pp. 896–912, May 2009.
[38] M. Portoles-Comeras, A. Cabellos-Aparicio, J. Mangues-Bafalluy, A. Banchs, and J. Domingo-Pascual, "Impact of transient CSMA/CA access delays on active bandwidth measurements," in *Proc. ACM IMC*, Nov. 2009, pp. 397–409.
[39] V. Ribeiro, R. Riedi, R. Baraniuk, J. Navratil, and L. Cottrell, "PathChirp: Efficient available bandwidth estimation for network paths," in *Proc. PAM*, Apr. 2003.
[40] J. Ridoux, D. Veitch, and T. Broomhead, "The case for feed-forward clock synchronization," *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 231–242, Feb. 2012.
[41] H. She, Z. Lu, A. Jantsch, and L.-R. Zheng, "Estimation of statistical bandwidth through backlog measurement," in *Proc. WoNeCa*, Mar. 2012, pp. 11–14.
[42] J. Strauss, D. Katabi, and F. Kaashoek, "A measurement study of available bandwidth estimation tools," in *Proc. ACM IMC*, Oct. 2003, pp. 39–44.
[43] B. White, J. Lepreau, L. Stoller, R. Ricci, S. Guruprasad, M. Newbold, M. Hibler, C. Barb, and A. Joglekar, "An integrated experimental environment for distributed systems and networks," in *Proc. OSDI*, Dec. 2002, pp. 255–270.

Ralf Lübben (S'11) received the Master's degree in computer science from the Leibniz Universität Hannover, Hannover, Germany, in 2007, and is currently pursuing the Ph.D. degree in communications technology at the Leibniz Universität Hannover.

Markus Fidler (M'04–SM'08) received the doctoral degree in computer engineering from RWTH Aachen University, Aachen, Germany, in 2004.

He was a Post-Doctoral Fellow with the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 2005, and the University of Toronto, Toronto, ON, Canada, in 2006. During 2007 and 2008, he was an Emmy Noether Research Group Leader with Technische Universität Darmstadt, Darmstadt, Germany. Since 2009, he has been a Professor of communications networks with Leibniz Universität Hannover, Hannover, Germany.

Jörg Liebeherr (S'88–M'92–SM'03–F'08) received the Ph.D. degree in computer science from the Georgia Institute of Technology, Atlanta, GA, USA, in 1991.

He was on the faculty of the Department of Computer Science with the University of Virginia, Charlottesville, VA, USA, from 1992 to 2005. Since 2005, he has been with the University of Toronto, Toronto, ON, Canada, as a Professor of electrical and computer engineering and the Nortel Chair of Network Architecture and Services.